

CadiDF una Herramienta de Soporte para la enseñanza-aprendizaje del Proceso de Normalización de Bases de Datos

Nadia Ramos¹, Rodolfo Bertone², Pablo Thomas², Germán Tejero¹,

¹ Universidad Nacional de Tierra del Fuego, Instituto de Desarrollo Económico e Innovación, Hipólito Irigoyen 880, 9410 Ushuaia, Tierra del Fuego, Argentina

² Universidad Nacional de La Plata, Facultad de Informática, Instituto de Investigación en Informática LIDI (III LIDI), 50 y 115 1er Piso, 1900 La Plata, Buenos Aires, Argentina

{nramos, ctejero}@untdf.edu.ar, {pbertone, pthomas}@lidi.info.unlp.edu.ar

Resumen. Las tecnologías informáticas y de la comunicación son utilizadas para la construcción del aprendizaje y el perfeccionamiento de habilidades, redefinen el rol del docente y de los estudiantes, mejorando todo el proceso educativo.

En las carreras de la disciplina Informática, no puede concebirse la enseñanza sin la transversalidad de estas tecnologías. En los últimos años, siguiendo una tendencia que se aplica en todos los ámbitos de la educación, los cursos vinculados a Bases de Datos buscan apoyarse en el uso de TICs a fin de favorecer los procesos de enseñanza y de aprendizaje.

En la asignatura introductoria a las Bases de Datos que se dicta en la Universidad Nacional de Tierra del Fuego (UNTdF) se desarrollan temas entre los que están presentes las formas normales y el proceso de normalización de relaciones basado en dependencias funcionales. Para esto, no se contaba con una herramienta de libre acceso, con una interfaz amigable, y que permita asistir a los alumnos en la identificación y aplicación de los algoritmos vinculados a su estudio.

En este trabajo se presenta el desarrollo de un prototipo “ad-hoc” llamado CadiDF, que permite trabajar sobre conjuntos de dependencias funcionales en una Base de Datos, la forma normal que se halla y su proceso de normalización.

Palabras clave: Enseñanza y aprendizaje de Bases de Datos, enseñanza y aprendizaje del proceso de normalización. Herramienta educativa. Formas normales. Bases de Datos.

1 Introducción

Las herramientas o plataformas informáticas y los diversos medios de comunicación masivos, pueden utilizarse como soporte para la mejora de los procesos de enseñanza-aprendizaje. Estas herramientas modifican el rol del docente, posicionándolo no sólo

como un gestor del conocimiento sino en el rol de creador del mismo, y en su vinculación con los alumnos [5].

El rol que desempeña el docente, como así también el alumno, el cambio en las fuentes de conocimiento y la didáctica, presuponen un nuevo paradigma en la educación. Este paradigma está basado en la optimización del proceso [5][6].

En una carrera relacionada con los sistemas informáticos resulta difícil contemplar el rol del docente como un gestor del conocimiento sin el apoyo de una herramienta informática, sobre todo en una asignatura referente a Bases de Datos.

Comunicar el conocimiento hacia los alumnos no sólo se circunscribe al horario de clases. El alumno debe afianzarlo también fuera de este ámbito, por lo que muchas de las dudas se plantean fuera del mismo. Actualmente la mayoría de las personas tienen acceso tanto a una computadora como a Internet para poder obtener conocimiento. Pero la apropiación de ese conocimiento no es tarea fácil, ya sea por la dificultad en su interpretación o por la falta del marco adecuado.

La utilización de herramientas informáticas como asistentes en el proceso de enseñanza-aprendizaje brinda un apoyo de suma importancia en el proceso educativo, permitiendo al alumnado afianzar el conocimiento adquirido durante los horarios de clase.

Más específicamente, la herramienta presentada en este trabajo brinda un soporte a los alumnos en la práctica de los conceptos de normalización de relaciones en una Base de Datos (BD). El producto desarrollado brinda un soporte al rol docente fuera y dentro de los horarios de clases.

El resto de este artículo se organiza del siguiente modo; en el capítulo 2 se presentan las ventajas de la utilización de las herramientas informáticas en el proceso de normalización de BD en contraposición al método tradicional, en el capítulo 3 se analizan cuatro herramientas web disponibles en el mercado para la enseñanza del proceso de normalización de BD. Luego en los capítulos 4 y 5 se presenta un breve marco teórico y la realización del prototipo CadiDF, respectivamente. Finalmente se presentan las conclusiones y trabajo futuro.

2 Enseñanza de la normalización

El proceso de normalización de una Base de Datos es un tema que se encuentra en la currícula de cualquier curso referido a Bases de Datos Relacionales en las instituciones de educación superior.

Los métodos de enseñanza tradicionales se han utilizado para presentar el proceso de normalización de las Bases de Datos. Con la creciente incidencia de los medios informáticos y comunicacionales, se brinda a los alumnos recursos de aprendizaje alternativos y visualmente atractivos, que pueden mejorar el proceso de enseñanza-aprendizaje.

Para las nuevas generaciones de alumnos, el uso de tecnologías no sólo les proporciona ambientes más orientados a lo visual, sino que se responsabilizan de su

aprendizaje. Estos hechos, no hacen que el método de enseñanza a través de la utilización de las tecnologías reemplace a la metodología tradicional, sino que funciona como un complemento y como una extensión al rol docente.

Aquellos alumnos que todavía no han tenido un contacto práctico con BD les puede resultar complejo entender el concepto de anomalías. Puede ser comprensible la forma que estas anomalías impactan con el diseño de BD, pero no han tenido la posibilidad de experimentar los inconvenientes que se producen con modelos que las tengan. Es por ello que la enseñanza de la teoría de normalización de BD se hace bajo conceptos que son abstractos y difíciles de comprender.

En la enseñanza tradicional hay una interacción con el docente. Se utilizan recursos clásicos como proyector y pizarrón para la explicación del tema. Se presentan al alumno ejercicios que deberán resolver durante el horario de clases o en sus horarios de estudio. Luego en el resto de la práctica, los alumnos realizan ejercicios teniendo al docente disponible en caso de surgir inquietudes.

Cuando el alumno resuelve ejercicios fuera del horario de clase, las dudas deberán ser aplazadas hasta el siguiente encuentro presencial o consultadas mediante otra vía: internet o libros disponibles en biblioteca.

Una de las mayores ventajas del método tradicional descripto es la interacción entre el docente y el alumno. Pero la enseñanza asistida con herramientas informáticas permite el acceso del docente al ámbito extra académico del alumno extendiendo sus funciones más allá de los límites físicos del aula y los horarios de clase.

En el caso del proceso de normalización de BD, los métodos visuales impartidos por una herramienta de software permiten a los alumnos apropiarse de los conceptos vistos en clase. El alumno puede verificar la correctitud de sus ejercicios y mejorar el ritmo de aprendizaje.

Las herramientas utilizadas aportan flexibilidad, pero no debe ser un reemplazo del rol docente. El alumno puede perjudicarse al trabajar aislado de sus compañeros y del docente, es por ello que las herramientas deben ser complementarias.

Más allá del método de impartir conocimiento que se utilice, tradicional o asistido con herramientas informáticas, la enseñanza de temas vinculados a BD continúa siendo un pilar de las carreras informáticas, sobre todo el diseño adecuado de BD a través del proceso de normalización.

El concepto a formalizar a través del prototipo CadiDF presentado en este trabajo, no necesita la implantación de un “ecosistema complejo” que almacene y manipule los datos. Estos pueden ser almacenados temporalmente dentro del navegador y utilizados durante el tiempo que permanezca el alumno utilizando la aplicación.

Al descartar la necesidad de almacenamiento de datos en una BD propiamente dicha, ya no es imperioso desarrollar un software back-end que permita manipular esos datos y administrar los accesos. Por ello se utilizó como almacenamiento temporal el store proporcionado por el navegador, no sólo para la relación sino para los cálculos derivados de la interacción con esta herramienta.

Esta aplicación web es una Single Page Application (SPA) generada con la librería ReactJS, aplicando el patrón Flux, a través del uso de Redux para almacenar los datos ingresados por el usuario y los resultados de los algoritmos de normalización en el

store. Una arquitectura SPA está compuesta por código HTML, CSS y JavaScript que se carga en un navegador. Carga el código necesario en una única página, que se actualiza en forma dinámica, en la medida en que es requerida. [7]

Por las características propias de la aplicación, no necesita implementación del lado del servidor. Únicamente se utilizará un servidor HTTP para aceptar la petición de la página en que se encuentra la aplicación.

En CadiDF sólo es necesaria la interacción con el usuario, éste no busca persistir los ejercicios ingresados. Se busca resolver el ejercicio en el momento. Esto hace que no haya separación entre front-end y back-end. Simplemente se elimina el back-end y se trabaja la resolución del problema del usuario en el cliente, es decir en el navegador del usuario.

3 Herramientas para la enseñanza del proceso de normalización

Si bien existe una gran cantidad de herramientas para la normalización de Bases de Datos, este proyecto tiene como propósito la creación de un software propio, orientado a mejorar el proceso de enseñanza aprendizaje.

En el desarrollo realizado se describen los conceptos básicos del proceso de normalización a partir de un conjunto de dependencias funcionales. El objetivo consiste en analizar las dependencias entre los atributos y la partición de entidades. No se busca obtener diseños correctos a partir de un diagrama de entidad interrelación.

Para el presente trabajo se analizaron algunas herramientas que son gratuitas y orientadas a la enseñanza, pero sólo se limitará a una interfaz web.

Las herramientas analizadas, fueron:

- Griffith University – Normalization Tool
- Tool for Database Design
- Relational Database Tools
- LDBN – Learn Database Normalization

No se evaluaron criterios de desarrollo o de cálculo, sino la forma en que se distribuye la información y cómo se muestra al usuario. Asimismo una característica importante es que la interfaz de la herramienta se vea en una forma amigable en los dispositivos móviles, esto considerando la masividad de estos dispositivos y su portabilidad.

El resultado del análisis de las cuatro herramientas mencionadas, ha considerado características que algunas de ellas poseen y funcionalidades que no incluyen, las cuales han sido incorporadas al prototipo desarrollado. Las características analizadas se presentan en la *Tabla 1*, a través de un cuadro comparativo.

Característica	Herramientas			
	Griffith University – Normalization Tool	Tool for Database Design	Relational Database Tools	LDBN – Learn Database Normalization
Carga intuitiva de atributos y dependencias	SI	NO	NO	SI
Repositorio de ejercicios	SI	NO	NO	SI
Descarga de ejercicio	SI	NO	NO	NO
Explicación de proceso	SI	NO	SI	NO
Incluye hasta Forma Normal Boyce Codd	SI	NO	SI	SI
Permite saber en que Forma Normal esta la relación	SI	NO	SI	NO
Interface amigable	SI	NO	NO	NO
Idioma Español	NO	NO	NO	NO
Diseño adaptable a móvil y tablet	SI	NO	NO	NO
Código disponible	NO	NO	SI	SI

4 Marco teórico para el desarrollo de CadiDF

El modelo relacional de BD plantea un conjunto de relaciones que representan la información del mundo real. Su representación se realiza a través de las relaciones y permite la manipulación de los datos en forma de actualización y consulta.

Un esquema de relaciones correctamente diseñado, no presenta anomalías de actualización. Las anomalías de actualización pueden encontrarse como repetición de datos o inconsistencias entre datos que debieran ser iguales. Las operaciones de eliminación y modificación de datos consumen mayor tiempo por necesitar la búsqueda de cada repetición de los datos a modificar. Además el agrupamiento incorrecto de atributos puede generar pérdida de información.

Las formas normales, surgen para identificar la presencia de anomalías en las relaciones. Cada relación está en una determinada Forma Normal (FN) siempre que cumpla con determinados requisitos. El proceso de normalización está basado en el análisis del esquema de relaciones de la BD, las dependencias funcionales que existen entre sus atributos y la identificación de atributos especiales denominados claves.

Existen siete (7) formas normales (1NF, 2NF, 3NF, BCNF, 4NF, 5NF y DKNF). Las primeras están basadas en el concepto de dependencia funcional. La Forma Normal de Boyce-Codd (BCNF) es una versión más restrictiva a la Tercera Forma Normal (3NF) como solución a las anomalías de actualización que pueden presentar estos esquemas bajo condiciones específicas. Finalmente se definieron tres formas normales adicionales que solucionan las redundancias y anomalías que pudieran

presentarse en esquemas en BCNF, pero no están basadas en el concepto de dependencia funcional, estas son: Cuarta Forma Normal (4NF), Quinta Forma Normal (5NF) y Forma Normal Dominio-Clave (DKNF).

En un esquema de Base de Datos se pueden especificar diversas restricciones. Estas restricciones pueden ser de dominio, de clave, de integridad referencial. También existen otro tipo de restricciones de dependencias de datos, entre las cuales se encuentran las restricciones de dependencias funcionales y dependencias multivaluadas.

Las restricciones de dependencias funcionales suelen ser utilizadas para la detección de anomalías y su eliminación a través del proceso de normalización.

La teoría de normalización permite obtener estructuras de datos eficientes y evita las anomalías de actualización. Es un método formal que permite realizar un diseño flexible de relaciones, que permita extenderse al agregar nuevos atributos a las relaciones existentes y nuevas relaciones [9].

Partiendo de una relación, sea realizan sucesivas proyecciones en relaciones más pequeñas, de modo que su concatenación permite obtener la relación original.

El proceso de normalización está basado en el análisis del esquema de las relaciones, en las dependencias funcionales que surgen de dicho esquema y sus claves candidatas. Cada esquema de relación que no cumple con los preceptos de la forma normal deseada se descompone en esquemas más reducidos que sí los cumplen.

El diseño de una Base de Datos con calidad puede realizarse siguiendo alguna de las siguientes metodologías: ascendente o descendente, o también llamadas metodologías de diseño por síntesis o análisis, respectivamente.

Ambos métodos persiguen el mismo fin, están basados en el concepto de la dependencia funcional y el cálculo del recubrimiento irredundante, pero tienen diferencias entre sí. La principal diferencia radica en que el método de análisis permite obtener una relación normalizada hasta Quinta Forma Normal (5NF), en cambio el método de síntesis permite llegar a Tercera Forma Normal (3NF) sin pérdida de información, y sino hubieran varias claves candidatas y solapadas podría llegarse a Forma Normal de Boyce-Codd (BCNF). [9][13]

5 CadiDF: una herramienta de soporte para los procesos de enseñanza y de aprendizaje de la práctica del proceso de normalización

En el contexto descripto, surge la realización de un prototipo que permitiera reemplazar el rol del papel y el lápiz en la práctica del proceso de normalización de relaciones.

En este prototipo se busca integrar las características destacadas de las herramientas web disponibles actualmente en el mercado, y los comportamientos observados durante las clases prácticas de la asignatura Bases de Datos I en la Universidad Nacional de Tierra del Fuego.

Para desarrollar este prototipo se analizó el comportamiento del alumnado durante cuatro (4) ciclos lectivos. Esto permitió observar qué dispositivos utilizan durante la clase, sus técnicas de estudio, las consultas que realizan sobre el proceso de normalización de BD y su trabajo áulico.

El proceso de normalización puede realizarse siguiendo el método de análisis o de síntesis. Ambas metodologías requieren de una relación, el cálculo del recubrimiento irredundante, las claves candidatas y la forma normal en la que se encuentra esta relación. Partiendo de estas componentes, la relación universal se proyecta en relaciones más pequeñas que satisfacen formas normales más avanzadas que la forma normal en la que se encuentra la relación.

Finalmente se obtienen n relaciones en la forma normal deseada, que concatenándose representan una relación equivalente a la relación universal. Esto permite conservar la información contenida en la relación (atributos y tuplas) y las dependencias [9][13].

CadiDF discrimina el cálculo del proceso de normalización en cuatro módulos: la carga del conjunto de atributos y dependencias funcionales que conforman la relación; el módulo calculadora, donde se presenta el resultado del cálculo de los conjuntos de relaciones del proceso de normalización; el módulo tutorial, detalla cada uno de los resultados obtenidos en el módulo calculadora con sus conjuntos intermedios; y el módulo desafío que permite al alumno ingresar las respuestas que considera solución y verificar su correctitud.

5.1 Módulo calculadora y tutorial

Para utilizar los módulos de CadiDF, previamente debe ingresarse una relación en su pantalla principal. Esta pantalla permite el ingreso de un conjunto de atributos y las dependencias funcionales entre ellos, como se muestra en la Fig. 1.

The screenshot shows the CadiDF interface with a dark header. The header contains the logo 'A-B CadiDF' and the tagline 'Aprende de una forma interactiva'. Below the header is a navigation bar with four icons: 'Nuevo problema', 'Calculadora', 'Tutorial', and 'Desafío'. The main content area is divided into two panels. The left panel, titled 'Atributos', has a text input field 'Ingresá el atributo...' with a green '+' button and a 'Limpiar todo' button. Below this are five rows of attribute input fields labeled A, B, C, D, and E, each with a green checkmark icon and a red 'X' icon. The right panel, titled 'Dependencias funcionales', has a text input field 'Implicante' followed by an arrow and a text input field 'Implicado' with a green '+' button and a 'Limpiar todo' button. Below this are five rows of functional dependency input fields: 'A → B', 'BCD → A', 'C → B', 'AD → D', and 'DE → B', each with a green checkmark icon and a red 'X' icon.

Fig. 1. Pantalla principal de CadiDF – Módulo de carga

Los módulos calculadora y tutorial están basados en los mismos algoritmos: calculan el recubrimiento irredundante, las claves candidatas y la forma normal en que se encuentra la relación para entregar un conjunto de relaciones en la forma normal deseada. En la Fig. 2 se muestra el cálculo de la clave candidata en el módulo tutorial, donde se explica su procedimiento.

CadiDF
Aprendé de una forma interactiva

Nuevo problema Calculadora **Tutorial** Desafío

Claves candidatas

1. Calcular el recubrimiento minimal
 - $A \rightarrow B$
 - $CD \rightarrow A$
 - $C \rightarrow B$
 - $DE \rightarrow B$
2. Del recubrimiento minimal, calcular el cierre transitivo de todos los determinantes
 - $A^+ = AB$
 - $CD^+ = CDAB$
 - $C^+ = CB$
 - $DE^+ = DEB$
3. Verificar si hay atributos que sólo sean implicantes en el conjunto de dependencias funcionales
Forman parte de todas las claves candidatas
4. Verificar si hay atributos que no participan en ninguna dependencia funcional
No hay
5. Verificamos si la unión de los atributos que no participan y los que participan sólo como implicantes en las dependencias funcionales es superclave
 - CDE
Como es superclave, entonces es la única clave candidata

C D E

Fig. 2. Procedimiento de cálculo del conjunto de claves candidatas en CadiDF –

Módulo tutorial

La posibilidad de seleccionar qué método se utilizará para normalizar la relación significó la implementación del algoritmo para dar soporte al método de análisis, realizando un proceso iterativo de descomposición de relaciones hasta lograr la forma normal deseada, y el algoritmo de síntesis que realiza un proceso de agrupación de atributos hasta llegar a tercera forma normal (3NF).

Los algoritmos para la resolución del problema, fueron seleccionados en concordancia con los contenidos teóricos de la asignatura. Esto permite a los alumnos entender el proceso que están estudiando e incorporarlo en una forma más visual.

5.1.1 Cierre de un descriptor

CadiDF utiliza el algoritmo tradicional de cálculo del cierre de un descriptor, que fuera propuesto por Ullman [15] y que recibe como entrada un conjunto de dependencias funcionales y un descriptor X .

5.1.2 Recubrimiento irredundante

El algoritmo de cálculo del recubrimiento irredundante es el que fuera propuesto por Ullman [15] y otros autores, que transforma un conjunto de dependencias funcionales elementales en un conjunto mínimo y equivalente. Este algoritmo está basado en el cálculo del cierre de un descriptor.

5.1.3 Claves candidatas

CadiDF utiliza los preceptos propuestos por Delobel para el cálculo de todas las claves candidatas de una relación [13]. Este algoritmo no utiliza el conjunto de dependencias funcionales ingresado por el usuario, sino que su entrada es el recubrimiento irredundante calculado previamente.

5.1.4 Forma normal de la relación

Para determinar la forma normal de la relación, se realizan verificaciones en orden creciente de forma normal, primero se verifica la segunda forma normal (2FN), luego la tercera forma normal (3FN) y por último la forma normal de Boyce-Codd (BCNF). No se realiza verificación alguna sobre la 1FN, dado que no pueden ingresarse atributos multivaluados en CadiDF.

5.2 Módulo desafío

Este módulo permite al alumno ingresar la solución propuesta por él y verificar su correctitud.

En virtud a que el recubrimiento irredundante de una relación puede no ser único, se calcula el cierre del conjunto de dependencias para corroborar la equivalencia con el conjunto mínimo de dependencias, y los demás algoritmos son calculados basándose en el recubrimiento ingresado.

6 Conclusiones y trabajos futuros

CadiDF proporciona a los alumnos una herramienta práctica para afianzar los conceptos teóricos concernientes al Proceso de Normalización de Bases de Datos. Permite corroborar la correctitud en las soluciones planteadas a los ejercicios prácticos, así como la descripción de los algoritmos para llegar a esas soluciones.

Este prototipo fue concebido como un recurso motivacional para el alumnado, transformándolo en hacedor de su propio aprendizaje, situándolo en una posición crítica hacia ellos mismos en la resolución de problemas.

La posibilidad del contacto con la práctica de normalización de relaciones, de estudiar el comportamiento del alumnado frente a esta disciplina han permitido y han guiado el desarrollo de CadiDF. La utilización durante el período de clases durante el segundo semestre 2018 permitirá, a través de la interacción con usuarios reales, una realimentación para la corrección de defectos, ampliar sus funcionalidades o mejorar el diseño e interoperabilidad.

Si bien CadiDF posee un diseño web responsivo, la utilización de React y Redux para su implementación permite que solo haya que reescribir partes para lograr una aplicación móvil, con React Native como librería de desarrollo en vez de React. Esto posibilita, como trabajo futuro, la conversión de CadiDF, a una aplicación móvil nativa.

Asimismo, como trabajo futuro se ha planteado a esta herramienta como un módulo de una herramienta más grande e integradora de más unidades que se dictan en la cátedra Bases de Datos I. Esto permitirá que el alumno pueda centrarse en los contenidos propios de la asignatura, en lugar de aprender a utilizar distintas herramientas.

Referencias

1. C. Domínguez Pérez, A.J. Elizondo y T.A. Pérez Fernández. "¿Todavía interesa normalizar bases de datos? Reflexionando sobre su enseñanza". XV JENUI, Barcelona, julio 2009
2. L. Ahmedi y E. Jajaga. "A database normalization tool using Semantic Web technologies", vol. 5, núm. 4, 2011, International Journal of Systems Applications, Engineering & Development
3. C. Stefanidis y G. Koloniari, (2016) "An interactive tool for teaching and learning database normalization". In Proceedings of the 20th Pan-Hellenic Conference on Informatics (PCI '16). ACM, Nueva York, artículo 18. Disponible en: <https://doi.org/10.1145/3003733.3003790>
4. A. Yazici y Z. Karakaya, (2007) "JMathNorm: A Database Normalization Tool Using Mathematica", International conference; 7th, Computational science, Beijing
5. UNESCO. (2010). Experiencias de formación docente utilizando tecnologías de información y comunicación. Santiago de Chile: AMF Imprenta.
6. D. Fernández Zalazar, et al. (2012). Los estudiantes universitarios y las TIC. de Moebius@cneri. Disponible en: web: <http://enmoebius.com.ar/?p=1846>
7. <https://reactjs.org/>

8. M.D. Costal Costa. El modelo relacional y el álgebra relacional. UOC/Digitalia: Universitat Oberta de Catalunya, 2001, pp. 7-31
9. R. Elmarsy y S.B. Navathe. Fundamentos de Sistemas de Bases de Datos. 5ta Ed. Madrid: Pearson - Addison Wesley, 2007
10. C.M. Ricardo. Bases de Datos. México: Mc Graw Hill, 2009, pp.123-198
11. E.F. Codd. A relational model of data for large shared data banks. IBM Research Laboratory, San José, California, Commun. ACM 13 (junio 1970)
12. J. Sánchez. (2013). Principios sobre Bases de Datos Relacionales (versión 2.4) [Online]. Disponible: <http://jorgesanchez.net/manuales/viejos/gbd/gbd2012.pdf>
13. A. de Miguel Castaño, M. Piattini Velthuis y E.M. Martínez. Diseño de Bases de Datos Relacionales. México: Alfaomega, 2000
14. R.C. Paré et al. Bases de Datos. Barcelona: fundació per a la Universitat Oberta de Catalunya, 2005
15. J.D Ullman y J. Widom. Introducción a los Sistemas de Bases de Datos. México: Prentice Hall, 1999
16. P. A. Bernstein. (1976). Synthesizing Third Normal Form Relations from Functional Dependencies.